The chemosensory transcriptome of a diving beetle

Nicolas Montagné, Muriel Jager, Thomas Chertemps, Emma Persyn, Yan Jaszczyszyn, Camille Meslin, Emmanuelle Jacquin-Joly, Michaël Manuel

▶ **To cite this version:**

Nicolas Montagné, Muriel Jager, Thomas Chertemps, Emma Persyn, Yan Jaszczyszyn, et al.. The chemosensory transcriptome of a diving beetle. Frontiers in Ecology and Evolution, 2021, 9, pp.773915. 10.3389/fevo.2021.773915 . hal-03510084

HAL Id: hal-03510084

https://hal.sorbonne-universite.fr/hal-03510084

Submitted on 4 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 03 December 2021
doi: 10.3389/fevo.2021.773915

The Chemosensory Transcriptome of a Diving Beetle

Nicolas Montagné[1]†, Muriel Jager[2†], Thomas Chertemps[1], Emma Persyn[2], Yan Jaszczyszyn[3], Camille Meslin[1], Emmanuelle Jacquin-Joly[1] and Michaël Manuel[2]**

[1] Sorbonne Université, INRAE, CNRS, IRD, UPEC, Université de Paris, Institute of Ecology and Environmental Sciences of Paris (iEES-Paris), Paris, France, [2] Sorbonne Université, MNHN, CNRS, EPHE, Institut de Systématique, Evolution, Biodiversité (ISYEB UMR 7205), Paris, France, [3] Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Gif-sur-Yvette, France

OPEN ACCESS

Edited by:
Rui Tang,
Institute of Zoology, Guangdong Academy of Science, Chinese Academy of Sciences (CAS), China

Reviewed by:
Hao Guo,
Chinese Academy of Sciences (CAS), China
Zhongzhen Wu,
Zhongkai University of Agriculture and Engineering, China

***Correspondence:**
Nicolas Montagné
nicolas.montagne@sorbonne-universite.fr
Michaël Manuel
michael.manuel@upmc.fr

† These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Chemical Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: 10 September 2021
Accepted: 17 November 2021
Published: 03 December 2021

Citation:
Montagné N, Jager M, Chertemps T, Persyn E, Jaszczyszyn Y, Meslin C, Jacquin-Joly E and Manuel M (2021) The Chemosensory Transcriptome of a Diving Beetle. Front. Ecol. Evol. 9:773915. doi: 10.3389/fevo.2021.773915

Insects astoundingly dominate Earth's land ecosystems and have a huge impact on human life. Almost every aspect of their life relies upon their highly efficient and adaptable chemosensory system. In the air, most chemical signals that are detected at long range are hydrophobic molecules, which insects detect using proteins encoded by multigenic families that emerged following land colonization by insect ancestors, namely the odorant-binding proteins (OBPs) and the odorant receptors (ORs). However, land-to-freshwater transitions occurred in many lineages within the insect tree of life. Whether chemosensory gene repertoires of aquatic insects remained essentially unchanged or underwent more or less drastic modifications to cope with physico-chemical constraints associated with life underwater remains virtually unknown. To address this issue, we sequenced and analyzed the transcriptome of chemosensory organs of the diving beetle *Rhantus suturalis* (Coleoptera, Dytiscidae). A reference transcriptome was assembled *de novo* using reads from five RNA-seq libraries (male and female antennae, male and female palps, and wing muscle). It contained 47,570 non-redundant unigenes encoding proteins of more than 50 amino acids. Within this reference transcriptome, we annotated sequences coding 53 OBPs, 48 ORs, 73 gustatory receptors (GRs), and 53 ionotropic receptors (IRs). Phylogenetic analyses notably revealed a large OBP gene expansion (35 paralogs in *R. suturalis*) as well as a more modest OR gene expansion (9 paralogs in *R. suturalis*) that may be specific to diving beetles. Interestingly, these duplicated genes tend to be expressed in palps rather than in antennae, suggesting a possible adaptation with respect to the land-to-water transition. This work provides a strong basis for further evolutionary and functional studies that will elucidate how insect chemosensory systems adapted to life underwater.

Keywords: ecological transitions, freshwater insects, Dytiscidae, chemical ecology, chemosensory receptors, odorant-binding proteins

INTRODUCTION

Chemical senses are at the crossroad between an animal and its environment, and thus play a key role in species adaptation (Yohe and Brand, 2018). Over hundreds of millions of years, insects have remained tremendously diverse and ecologically successful in every kind of continental ecosystem on Earth. This owes a great deal to the characteristics and evolvability of their chemosensory system,

finely tuned to detect chemical cues in an aerial environment. Although the vast majority of insect species are terrestrial, there are also many insects that live in freshwater habitats, where they are abundant and diversified worldwide (Dijkstra et al., 2014). All are members of lineages derived from terrestrial ancestors, which have secondarily adapted to life in water. This raises the question of how the fundamentally aerial chemosensory equipment of insects has been remodeled in these lineages, to cope with the drastically different physico-chemical constraints associated with an aquatic environment. Indeed, the compounds most readily displaced over long distances in water are predominantly hydrophilic whereas those displaced in the air are mainly hydrophobic (Mollo et al., 2014). Until recently, it was thought that aquatic animals were only able to detect hydrophilic molecules. However, hydrophobic compounds are also prevalent in aquatic ecosystems, and it has recently been shown that marine shrimps and freshwater fish can perceive hydrophobic molecules, either by contact or at short-range (Giordano et al., 2017).

In aerial insects, olfactory sensory neurons—involved in long-range chemodetection—are mainly localized on antennae but can also be found on other body parts, including maxillary and labial palps, depending on the taxa (Hansson and Stensmyr, 2011). Gustatory sensory neurons—involved in short-range or contact chemodetection—are found on palps and legs, as well as antennae, wings and ovipositors (Montell, 2009). Detection of odorants and tastants by these peripheral neurons is mediated by different families of chemoreceptor proteins localized in their dendritic membrane. Among these, an insect-specific family of chemoreceptors called odorant receptors (ORs) arose after water-to-land transition (ca. 410 My ago), in a common ancestor of Ectognatha (Brand et al., 2018). ORs are expressed in olfactory neurons and bind airborne hydrophobic molecules (e.g., terpenoids, benzenoids, fatty acid derivatives, . . .). They are transmembrane proteins associated in heteromers with a unique co-receptor named Orco, forming a non-selective cation channel that opens upon ligand binding (Wicher and Miazzi, 2021). In addition to ORs, two major chemoreceptor families have been characterized through genomic and functional studies in aerial insects. The gustatory receptors (GRs) belong to the same superfamily as ORs, though they are not specific to insects (Eyun et al., 2017; Robertson, 2019). They are expressed in gustatory sensory neurons and some have been shown to bind CO_2, sugars or bitter molecules (Isono and Morita, 2010; Robertson, 2019; Xu, 2020). The third family of chemosensory receptors in insects are the ionotropic receptors (IRs), also found in other protostomes (Croset et al., 2010). To date, the function of IRs has been studied mainly in *Drosophila*, where they are expressed in olfactory, gustatory and other sensory neurons and form complexes with several co-receptors (Silbering et al., 2011; Sanchez-Alcaniz et al., 2018). Antennal IRs are primarily responsible for the detection of volatile hydrophilic molecules such as acids and amines, but some are also involved in temperature or humidity sensing (van Giesen and Garrity, 2017; Rimal and Lee, 2018). Outside of *Drosophila*, involvement of antennal IRs in odorant detection has also been demonstrated in Hymenoptera and Lepidoptera (Shan et al., 2019; Zhang et al., 2019, 2021). Whatever the family considered, the size of chemoreceptor repertoires varies widely among insects, from ten to several hundreds of genes per species. This variability reflects their rapid evolution following a birth-and-death process, characterized by numerous gene duplications and losses (Robertson, 2019).

In addition to receptors, other chemosensory gene families are involved in the detection of chemical cues in insects. This is notably the case of odorant-binding proteins (OBPs), secreted in high concentration in the lymph of olfactory sensilla, where they solubilize hydrophobic molecules to facilitate their transport to the chemoreceptors. These proteins thus contribute to the sensitivity of the olfactory system in aerial insects (Brito et al., 2016; Rihani et al., 2021). In addition to OBPs, which are insect-specific, other soluble proteins such as chemosensory proteins (CSPs) and Niemann-Pick C2 (NPC2) proteins are suspected to play a role in chemical senses (Pelosi et al., 2014). The lymph of olfactory sensilla also contains large amounts of odorant-degrading enzymes, belonging to various enzymatic families, important for rapid signal termination following receptor activation (Leal, 2013; Chertemps and Maïbèche, 2021).

The evolution of chemosensory genes has likely played a major role in adaptation of insects to freshwater habitats. However, we know virtually nothing about chemosensory genes in aquatic insects besides pioneering works in mosquito larvae (Xia et al., 2008; Liu et al., 2010; Ruel et al., 2018). In this study, we aimed to gain insights into evolutionary changes that affected chemosensory gene repertoires after land-to-water transition, through transcriptome analysis of chemosensory organs of a diving beetle (Coleoptera, Dytiscidae). Among the many Coleoptera lineages in which such transitions occurred (Short, 2018), Dytiscidae exhibit the highest degree of adaptation to an aquatic life. Adult diving beetles are capable of flying from one water body to another, but they spend most of their life in water and notably feed and reproduce only underwater. They are predaceous animals, like most other members of the Coleoptera suborder Adephaga, and chemical senses may play a prominent role in prey detection (Culler et al., 2014). Morphological studies have indicated that porous sensilla at the surface of antennae and palps of Dytiscidae differ from those of their closest terrestrial cousins, the Carabidae (Baker, 2001). In diving beetles, there is experimental evidence suggesting that the antennae play a role in chemoreception both underwater and in the air, whereas the palps detect chemical stimuli in the liquid medium only (Hodgson, 1953). Furthermore, active movements of the maxillary palps occur at rest in response to exposure to food odors, during subsequent swimming when attempting to locate preys, and upon feeding (M. Manuel, personal observations). However, very little is known concerning the physiology of these chemosensory structures in diving beetles.

Here we have sequenced, assembled and analyzed the transcriptomes of antennae and palps from adult males and females of the diving beetle species *Rhantus suturalis* Macleay, 1825. This species is of moderate size (10.5–12.5 mm), is common and widespread from western Europe and North Africa through Asia to northern Australia, and lives in a wide variety

of freshwater lentic habitats, with a preference for more or less temporary ponds in open environments. Species of the genus *Rhantus* are prominent predators of mosquito larvae (Culler et al., 2014). Furthermore, it was recently demonstrated that *R. suturalis* males are attracted underwater by a sex pheromone of unknown composition emitted by females (Herbst et al., 2011). The *R. suturalis* transcriptome was used to annotate genes belonging to major families of soluble proteins and transmembrane receptors responsible for semiochemical detection in insects, and to estimate their expression levels in antennae and palps. In parallel to this candidate gene approach, we also searched genes specifically expressed in these chemosensory tissues relative to wing muscle, and genes specifically expressed in one sex vs. the other. By doing so, our main goal was to identify features of the diving beetle chemosensory gene toolkit that may reflect specificities associated with an aquatic life, notably expansions or contractions of gene repertoires and unusual expression patterns in antennae and palps when compared to terrestrial insects.

MATERIALS AND METHODS

Sample Preparation and RNA Extraction

Adult males and females of *R. suturalis* were collected in several ponds in the Fontainebleau forest (Bois-le-Roi, ca. 48°28′N 2°39′E, France) and kept alive in the laboratory in small water tanks. Ablations of antennae and palps were performed under a Leica MZ16 stereomicroscope, with the specimen placed upside down and immobilized in a custom device. Appendages were removed by grasping with forceps at the most basal article (antennomere or palpomere I). Samples and numbers of individuals used for each sample (in brackets) were as follows: female antennae (50), male antennae (45), female palps (maxillary and labial palps mixed in the same sample, 79), male palps (69) and wing muscle (one male individual), leading to a total of five samples. Each specimen was killed in liquid nitrogen immediately after appendage ablations. Each appendage was rinsed in RNAse free water, then immediately transferred into a 2 ml tube on ice containing 500 µl TRIzol™ Reagent (Thermo Fisher Scientific, Waltham, MA) and a mixture of micro-ceramic beads. Tubes were frozen during at least one night at –20°C.

For RNA extraction, 300 µl of TRIzol™ Reagent was added to the tubes prior to grinding, lysis and homogenization (3 × 30 s, then 5 min on ice, then 3 × 30 s all cycles at 5,000 rpm) in a Minilys homogenizer (Bertin technologies SAS, Montigny Le Bretonneux, France). The liquid phase was then separated from the beads and total RNA was extracted using the phenol/chloroform method as described in the TRIzol™ Reagent user guide. Total RNA was treated with TURBO DNase (Thermo Fischer Scientific) according to manufacturer's instructions, then purified and concentrated using the RNeasy® MinElute™ Cleanup Kit (Qiagen, Hilden, Germany). RNA quality and quantity were measured using a ND-1,000 NanoDrop spectrophotometer (Thermo Fisher Scientific). All RNA samples were also analyzed on a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA) to determine their RNA Integrity Numbers (RIN). All measured RIN were > 7.

RNA Sequencing and *de novo* Transcriptome Assembly

cDNA library construction and sequencing were carried out at the high-throughput sequencing facility of the Institute for Integrative Biology of the Cell (Gif-sur-Yvette, France). Libraries have been prepared using the Illumina TruSeq mRNA Stranded kit, with minor modifications allowing to obtain long cDNA inserts. Sequencing (150-bp paired-end reads) was performed using a NextSeq500 instrument. Data processing and analysis (summarized in **Supplementary Figure 1**) was performed on the Galaxy server hosted by the BioInformatics Platform for Agro-ecosystems Arthropods (Rennes, France). Quality check was done with FastQC 0.69 and reads were trimmed with Trimmomatic 0.32 (Bolger et al., 2014). Parameters were as follow: Sliding window = 4 bases; average quality = 20; minlen = 30 bases; headcrop = 10 bases; trailing min. quality = 20. Clean reads from the five samples were assembled using Trinity 2.4 (Grabherr et al., 2011), with the following parameters: min contig length = 200, min count for k-mers to be assembled = 1. Coding sequences were extracted from the reference transcriptome using Transdecoder 3.0 (Haas et al., 2013) with a minimum protein length of 50 amino acids. Redundant sequences were then clustered using CD-HIT-EST 1.3 (Fu et al., 2012) with a similarity threshold of 0.9 and a word size of 8.

Transcriptome Analysis

Transcriptome quality was estimated using BUSCO 3.0 (Simao et al., 2015) with the insecta_odb9 dataset. For transcript annotation, coding sequences were first translated with Transeq 5.0 (Rice et al., 2000). Then, an alignment search was performed using DIAMOND (Buchfink et al., 2015) on the NCBI non-redundant (nr) protein sequence database (access March 2020) with the "more sensitive" mode, a BLOSUM62 scoring matrix and a maximum *e*-value of 1e-05. In parallel with the alignment search strategy, a protein domain analysis was performed with hmmscan 3.2 (Finn et al., 2011) using the Pfam-A hidden Markov model database (access May 2020; El-Gebali et al., 2019) and default parameters.

To measure expression levels, reads generated for each of the five libraries were aligned on the reference transcriptome with HISAT 2.1.0 (Kim et al., 2015). Transcript abundance was then measured in each sample as FPKM (fragments per kilobase of exon per million fragments mapped) with the Cufflinks 2.2.1 suite (Trapnell et al., 2012), with default parameters. Heatmaps were built using $\log_2$(FPKM + 1) values. Transcripts specifically expressed in chemosensory organs (antennae or palps) vs. wing muscle and transcripts specifically expressed in chemosensory organs of one sex vs. the other were identified using the Cuffdiff tool, with a False Discovery Rate of 0.05. Amino acid sequences translated from these transcripts were used as queries to search the UniProt database (access May 2020) using BLASTp (Johnson et al., 2008), with a *e*-value cutoff of 1e-03.

For the comparison of chemosensory gene expression bias in antennae and palps of *R. suturalis* and *Tribolium castaneum*, data used for *T. castaneum* were RPKM values provided in previously published transcriptomic analyses (Dippel et al., 2014, 2016). For each family (OBP, OR, GR, IR), orthologous relationships between *R. suturalis* and *T. castaneum* genes were determined using neighbor-joining phylogenies built with Seaview 4.7 (Gouy et al., 2010). Genes with FPKM or RPKM values < 1 in both tissues and genes with no orthologous relationship identified were discarded.

Annotation of Chemosensory Genes

For each gene family to be annotated, datasets containing amino acid sequences manually annotated in other coleopteran species were first created. The OR dataset consisted in sequences previously annotated (Mitchell et al., 2020) in the genomes of *Calosoma scrutator* (Carabidae), *Nicrophorus vespilloides* (Silphidae), *Agrilus planipennis* (Buprestidae), *Anoplophora glabripennis* (Cerambycidae), and *Dendroctonus ponderosae* (Curculionidae). The GR, IR, OBP, and CSP datasets contained sequences from *A. glabripennis* (McKenna et al., 2016), *A. planipennis*, and *D. ponderosae* (Andersson et al., 2019). The NPC2 dataset contained sequences from *T. castaneum* (Pelosi et al., 2014). These amino acid sequences were used as queries to search the *R. suturalis* reference transcriptome using tBLASTn 2.5 (Cock et al., 2015) with an *e*-value cutoff set at 1e-10. In parallel, results of the hmmscan analysis were mined for the following domains: pfam03392, OS-D Insect pheromone-binding family (CSPs); pfam01395, PBP/GOBP family (OBPs); pfam02949, 7tm_6 Odorant receptor (ORs); pfam08395, 7tm_7 Chemosensory receptor (GRs); pfam00060, Ligand-gated ion channel (IRs). To verify annotations and eliminate false positive hits, amino acid translations of the unigenes were searched against the NCBI nr database using BLASTp (Johnson et al., 2008). In some cases, redundant unigenes encoding the same protein but not clustered by CD-HIT-EST were manually clustered to rebuild a longer sequence. The presence of signal peptides within sequences of soluble protein precursors was predicted with SignalP 4.1 (Nielsen, 2017) and the presence of transmembrane domains within sequences of candidate chemoreceptors was predicted with TMHMM 2.0 (Krogh et al., 2001).

Phylogenetic Analyses

To rebuild the phylogenies of the CSP, OBP, OR, GR, and IR families, amino acid sequences from *R. suturalis* were aligned with the coleopteran sequences described above. Datasets were purged of predicted pseudogenes and isoforms resulting from alternative splicing. CSP, OBP, and IR alignments were performed with MAFFT 7 (Katoh et al., 2019) and OR and GR alignments were performed with Muscle (Edgar, 2004) as implemented in Seaview 4.7 (Gouy et al., 2010), then manually curated. The best model of amino acid substitution was determined by SMS (Lefort et al., 2017) and trees were calculated with the maximum-likelihood method using PhyML 3.0 (Guindon et al., 2010), with a mix of SPR and NNI algorithms. Node support was assessed with the SH-like approximate likelihood-ratio test (aLRT) as implemented in PhyML (Anisimova and Gascuel, 2006).

Quantitative Real-Time PCR

Relative gene expression levels were estimated by quantitative real-time PCR (qPCR) for 12 selected genes (3 genes for each of the OBP, OR, IR, and GR families) showing contrasted expression profiles in RNAseq results, in order to validate gene expression data derived from RNAseq. qPCR estimates were performed on RNA preparations independent from those used for RNAseq, and with maxillary and labial palps treated as two distinct samples. Twenty-five *R. suturalis* adults (11 males and 14 females) were collected in a pond in Rue (Somme, France) in October 2020 and kept alive in the lab for a few weeks before sample preparation. Dissections and RNA extractions were performed as described above, except that maxillary and labial palps were separated and that male and female tissues were mixed. First-strand cDNAs were synthesized using 500 ng of total RNA, oligo-dT primer and SuperScript™ II Reverse Transcriptase (Thermo Fischer Scientific). The pooled cDNA used to build standard curves for calculating PCR efficiencies was synthesized from a mix of 200 ng of RNA from antennae, 100 ng of RNA from each palp and 100 ng of RNA from wing muscle. Specific primer pairs were designed with OligoAnalyzer™ (Integrated DNA Technologies, Coralville, IA) to amplify fragments between 150 and 230 bp (primer sequences are given in **Supplementary Table 3**). These genes were selected based on the RNAseq data in order to have—for each family—one gene expressed in both antennae and palps, one antenna-biased gene and one palp-biased gene. The gene encoding the ribosomal protein RsutRPL13 was used as the reference gene.

qPCR assays were performed in technical duplicates on the pooled cDNA dilutions (undiluted, 1/5, 1/25, 1/50, 1/100, 1/200, 1/400), and in technical triplicates on the four cDNA samples (antennae, maxillary palps, labial palps, wing muscle). The qPCR mix contained 2 µl of cDNA (or water for the negative controls), 8 µl of SsoAdvanced™ universal SYBR® Green Supermix (Bio-Rad, Hercules, CA), and 500 nM of both gene-specific primers, in a final volume of 10 µl. PCR reactions were run in 96-well plates, in a CFX96 Touch™ Real-Time PCR detection system (Bio-Rad) with the following thermal cycling conditions: 98°C for 4 min; 40 amplification cycles at 98°C for 15 s, 60°C for 1 min. After amplification, melt curve analyses were performed by gradual heating from 65 to 95°C at 0.5°C.s^{-1}. Only one peak was detected for each sample. The slopes of the standard curves were calculated and the amplification efficiency was estimated as $E = (10^{-1/slope})$. Mean normalized expression of the target genes were calculated with Q-Gene (Simon, 2003).

RESULTS

The *Rhantus suturalis* Reference Transcriptome

Illumina sequencing of the five libraries generated a total of 526 million pairs of raw reads. After trimming, we finally obtained 72,793,754 pairs of clean reads for female antennae, 104,087,147

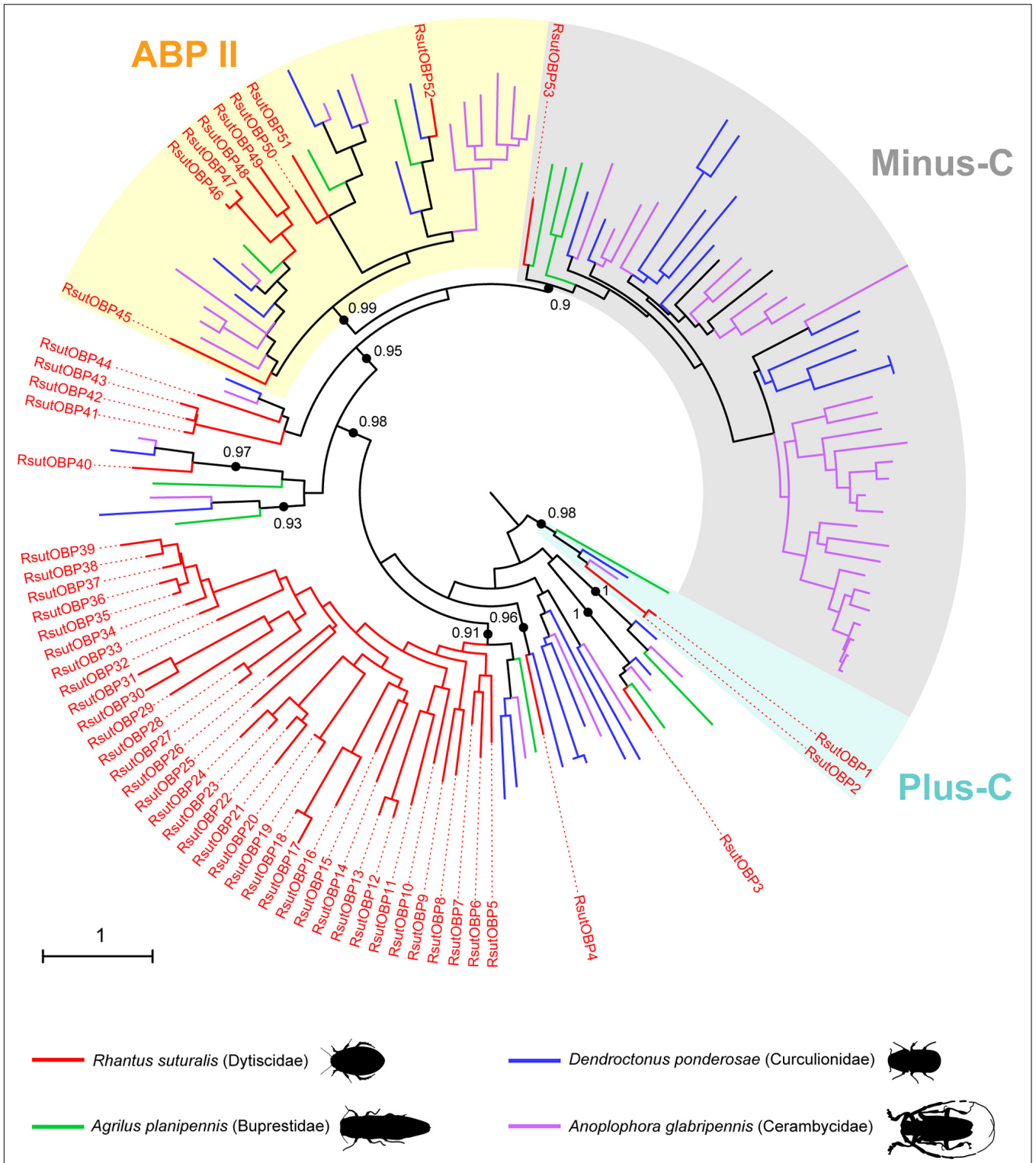

FIGURE 1 | Maximum-likelihood phylogeny of coleopteran OBPs. The tree was built from an alignment of amino acid sequences of *R. suturalis* OBPs (RsutOBP, available in **Supplementary Table 1**) with sequences annotated from the genomes of three coleopteran species belonging to distinct superfamilies. The Plus-C, Minus-C and Antennal Binding Protein II (ABPII) clades were defined as in Andersson et al. (2019). Black dots indicate deep nodes highly supported by the likelihood-ratio test (aLRT > 0.9), with the corresponding value. The tree was rooted using the Plus-C clade as an outgroup. The scale bar indicates the expected number of amino acid substitutions per site.

pairs for female maxillary and labial palps, 85,494,452 pairs for male antennae, 94,355,486 pairs for male palps and 86,825,522 pairs for the male wing muscle sample. These clean reads were pooled and assembled into ∼300,000 contigs, which led to a reference transcriptome of 47,570 unigenes encoding proteins of more than 50 amino acids (**Supplementary Figure 1**). The BUSCO analysis revealed a good completeness of this reference transcriptome (only 4.1% of missing genes) as well as a very low level of redundancy (2.7% of duplicated sequences). We found corresponding hits in the NCBI nr protein sequence database for 26,403 of the 47,750 unigenes, and identified conserved protein domains (from the Pfam database) for 27,553 unigenes.

Soluble Proteins Potentially Involved in Chemical Sensing

We annotated 53 transcripts encoding candidate OBPs (including 17 full-length coding sequences) in the reference transcriptome (**Supplementary Table 1**). Among these 53 RsutOBPs, we identified two members of the Plus-C clade (RsutOBP1-2), an OBP sub-family characterized by a number of cysteine residues above six (the number usually observed in OBPs), eight members of the Antennal Binding Protein II clade (RsutOBP45-52) and a single member of the Minus-C clade (RsutOBP53), characterized by the presence of only four cysteine residues instead of six (**Figure 1**). The other RsutOBPs clustered in clades generally referred to as "classical OBPs" (Dippel et al., 2014; Andersson et al., 2019). Among them, we identified a remarkable lineage-specific expansion, with RsutOBP5-39 (i.e., 35 out of the 53 OBPs) belonging to a single clade, which contained only two OBPs from *D. ponderosae* and one OBP from both *A. planipennis* and *A. glabripennis*.

Although the lack of biological replicates did not allow to unambiguously demonstrate differential gene expressions between tissues, we used FPKM values to estimate transcript abundance in each tissue. Moreover, quantitative real-time PCR on selected genes exhibiting contrasted expression patterns further confirmed the results obtained by RNAseq (**Supplementary Figure 2**). Reliable expression in chemosensory organs (i.e., antennae and palps) was observed for 48 of the 53 RsutOBP unigenes annotated, with no visible sexual dimorphism (**Figure 2** and **Supplementary Table 1**). Among these, we found 17 RsutOBPs exhibiting similar FPKM values in antennae and palps. This was notably the case for members of the Plus-C (RsutOBP1 and 2) and Minus-C (RsutOBP53) clades. Eight RsutOBPs appeared more expressed in antennae than palps (ratio of FPKM values above 4), all but one (RsutOBP45-52) belonging to the ABPII clade. Finally, 23 RsutOBPs seemed more expressed in palps than in antennae and most of these OBPs belong to the large RsutOBP expansion (RsutOBP5-39).

In addition to OBPs, we annotated four *R. suturalis* transcripts encoding CSPs and three transcripts encoding NPC2 proteins (**Supplementary Table 1**). RsutCSPs belonged to conserved lineages in the coleopteran CSP phylogeny (**Supplementary Figure 3**). Among the four candidate CSPs, two were highly expressed in chemosensory organs and therefore likely to be involved in chemical sensing: RsutCSP3 was expressed in antennae of both sexes whereas RsutCSP4 exhibited extremely high expression levels in both antennae and palps, as well as faint expression in muscle. None of the NPC2 proteins identified

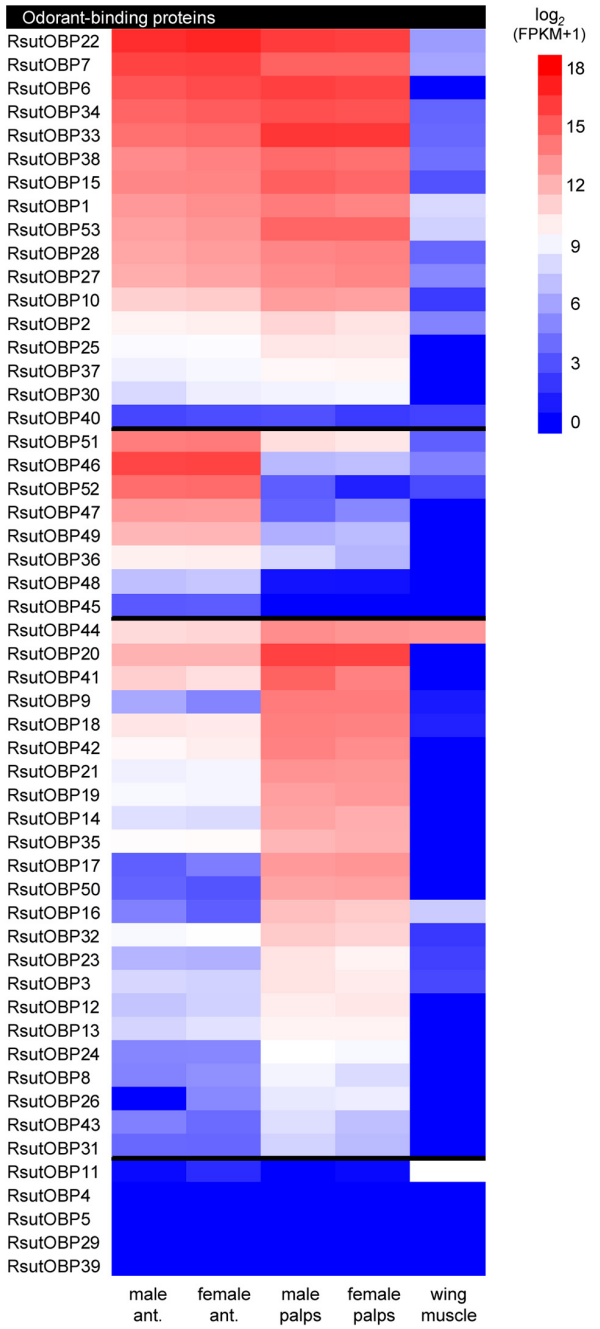

FIGURE 2 | Heatmap showing expression levels of *R. suturalis* transcripts encoding OBPs. Color coding is based on $\log_2$(FPKM + 1) values; raw FPKM values are available in **Supplementary Table 1**. Transcripts were classified in three categories, from top to bottom: equally expressed in antennae and palps; more expressed in antennae; more expressed in palps. A transcript was classified as more expressed in a specific tissue when the ratio of FPKM values was at least four-fold. Transcripts with FPKM values lower than 1 in all tissues are not shown.

was specific to chemosensory organs, with RsutNCP2_2 and 3 being moderately expressed in all tissues investigated (**Supplementary Table 1**).

Candidate Chemosensory Receptors

Altogether, we annotated transcripts encoding 48 ORs (30 full-length), 73 GRs (47 full-length) and 53 IRs (24 full-length) in the reference transcriptome (**Supplementary Table 1**). The *R. suturalis* ORs were distributed among six of the eight clades of the coleopteran OR phylogeny (**Figure 3**). Within each clade, RsutORs clustered with ORs annotated in the ground beetle *C. scrutator* (Carabidae), also belonging to the sub-order Adephaga. We found no RsutOR belonging to clades 5 and 6, but RsutOR9 and four *C. scrutator* ORs formed a clade that could be specific to Adephaga, with no Polyphaga representative. As ORs are the only chemosensory genes that have been annotated in *C. scrutator*, the OR tree is the only one in which duplications more recent than the divergence of the dytiscid lineage with respect to that of the terrestrial carabids can be identified. In addition to a few instances of such duplications that involve only two or three *R. suturalis* genes (e.g., RsutOR1–3 and RsutOR6 and 7, within clade 1), a significant OR expansion was apparent within clade 3, which gave rise to 9 *R. suturalis* paralogs (RsutOR26–34). As expected, the OR family member with the highest expression level was the obligate co-receptor RsutOrco, expressed at high levels in both antennae and palps (**Figure 4** and **Supplementary Table 1**). Most RsutORs (31 out of 48) were found reliably expressed only in male and female antennae, some with high expression levels. However, eight RsutORs appeared expressed in palps and not in antennae. All but one (RsutOR1) belong to the RsutOR expansion within clade 3 (RsutOR26–34). Interestingly, the two palpal RsutOR genes studied by qRT-PCR were found expressed in maxillary palps but not labial palps (**Supplementary Figure 2**).

In the GR phylogeny, RsutGR1–3 belong to the candidate CO_2 receptor clade, RsutGR4 and 5 to the candidate sugar receptor clade, and RsutGR10 to the candidate fructose receptor clade (**Figure 5**). The vast majority of GRs identified in *R. suturalis* (RsutGR15–70) belong to a single clade of the GR phylogeny without reported putative function. This clade contained only a few representatives from Polyphaga species and several occurrences of massive expansions of *R. suturalis* GRs. The vast majority of RsutGRs were either found expressed only in palps or more expressed in palps than in antennae (**Figure 4** and **Supplementary Table 1**). They generally exhibited low FPKM values, with the exception of the candidate CO_2 receptors RsutGR1–3 and a few others (RsutGR9, 36, 37, 72, and 73). The GR gene with the highest FPKM value—RsutGR9—was found specifically expressed in labial palps by qRT-PCR (**Supplementary Figure 2**).

Based on the IR phylogeny (**Figure 6**), we identified unigenes encoding each of the eight following "antennal" IRs, whose putative functions have been assigned based on previous work on the fruit fly (van Giesen and Garrity, 2017; Rimal and Lee, 2018): the four IR co-receptors IR25a, 8a, 93a, and 76b, the two humidity-sensing IR40a and 68a, the temperature-sensing IR21a and the olfactory IR41a. We also identified six RsutIRs within the IR75 clade (RsutIR75a-f), supposedly involved in the detection of water-soluble semiochemicals carrying a carboxylic acid or an amine function (Silbering et al., 2011; Prieto-Godino et al., 2017). The other 39 IRs found in the *R. suturalis* transcriptome belong to the divergent IR clade, involved in taste in *Drosophila* (Koh et al., 2014; Sanchez-Alcaniz et al., 2018). Several independent RsutIR expansions were found within this clade, including two large groups of 12 and 13 *R. suturalis* paralogs, respectively. RsutIR25a (candidate universal IR co-receptor) was the most highly expressed coreceptor, in both antennae and palps (**Figure 4** and **Supplementary Table 1**). RsutIR76b (candidate gustatory IR coreceptor) exhibited a high expression in palps, whereas RsutIR8a (candidate IR olfactory coreceptor) and RsutIR93a were highly expressed in antennae, like the four conserved antennal IRs (IR21a, 40a, 41a, 68a). Two RsutIRs belonging to the IR75 clade were also highly expressed in antennae, and three others exhibited similar FPKM values for both antennae and palps. Among these, RsutIR75f exhibited an exceptionally high expression level in both tissues (**Figure 4** and **Supplementary Figure 2**). The 39 divergent RsutIRs were globally expressed at low levels in palps, although a few of them were also expressed in antennae, sometimes specifically (e.g., RsutIR130). As observed for the soluble proteins, none of the candidate chemosensory receptors annotated here exhibited a sexually dimorphic expression.

Representatives of Other Protein Families Expressed in Chemosensory Organs

Complementary to the detailed analysis of gene families described above, we searched for unigenes from other families that would be specifically expressed in the chemosensory organs relative to the wing muscle. By doing so, we found more than 700 unigenes with hits in UniProt or Pfam databases (**Supplementary Table 2**). We notably identified more than 40 sequences encoding enzymes that may play a role in metabolism of semiochemicals, such as cytochrome P450 enzymes, carboxylesterases, lipases, short-chain dehydrogenases and UDP-glucuronosyl transferases (Chertemps and Maïbèche, 2021). Some of these unigenes exhibited high FPKM values (**Figure 7**). We also found six unigenes encoding CD36 proteins, a family to which Sensory Neuron Membrane Proteins (SNMP) belong (Zhao et al., 2020), and three unigenes encoding candidate pickpocket channels (amiloride-sensitive sodium channels), some of which are involved in olfaction and taste in *Drosophila* and mosquitoes (Chen et al., 2010; Matthews et al., 2019; Ng et al., 2019; Liu et al., 2020). We also searched for unigenes that would be expressed in chemosensory organs of a single sex (either male or female) and found only a dozen of such unigenes, i.e., five expressed in antennae and 10 in palps. None of these corresponded to gene families with a known direct or indirect role in chemical senses (**Supplementary Table 2**). This confirmed our initial observation that no chemosensory gene seemed differentially

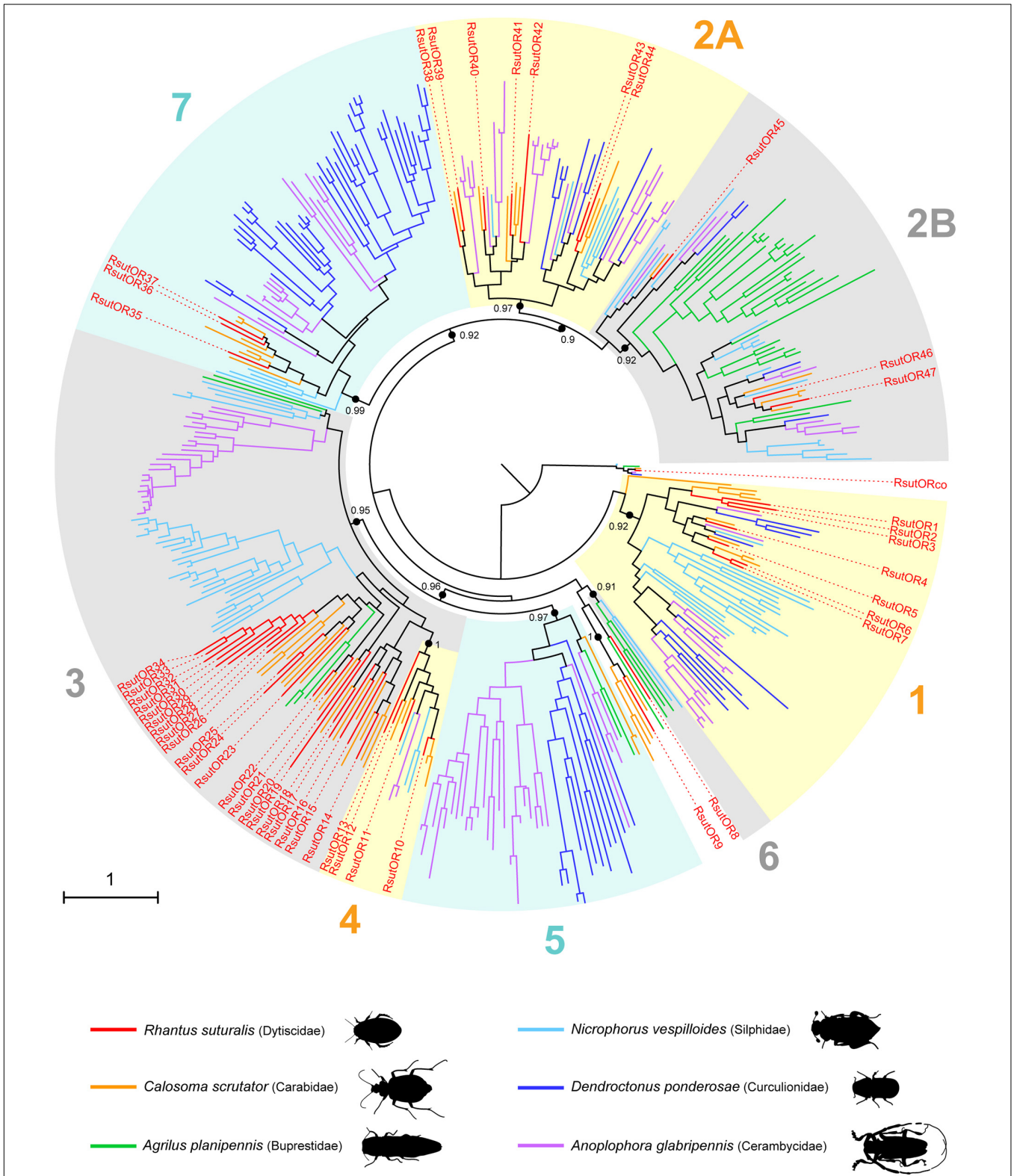

FIGURE 3 | Maximum-likelihood phylogeny of coleopteran ORs. The tree was built from an alignment of amino acid sequences of *R. suturalis* ORs (RsutOR, available in **Supplementary Table 1**) with sequences annotated from the genomes of five coleopteran species belonging to distinct superfamilies. The OR clades 1–7 were defined as in Mitchell et al. (2020). Black dots indicate deep nodes highly supported by the likelihood-ratio test (aLRT > 0.9), with the corresponding value. The tree was rooted using the Orco clade as an outgroup. The scale bar indicates the expected number of amino acid substitutions per site.

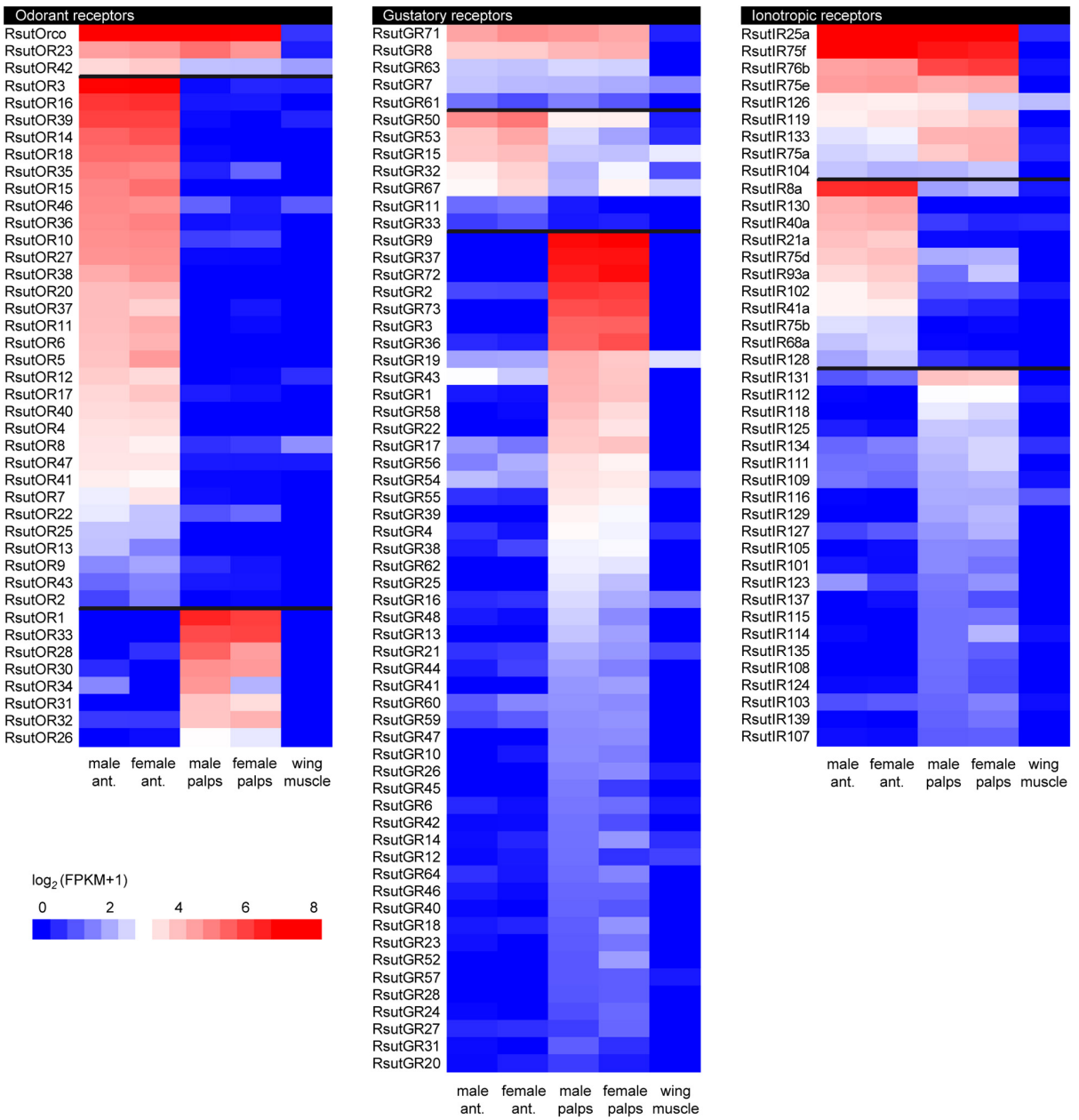

FIGURE 4 | Heatmaps showing expression levels of *R. suturalis* transcripts encoding candidate chemosensory transmembrane receptors. Color coding is based on $\log_2$(FPKM + 1) values; raw FPKM values are available in **Supplementary Table 1**. Transcripts were classified in three categories, from top to bottom: equally expressed in antennae and palps; more expressed in antennae; more expressed in palps. A transcript was classified as more expressed in a specific tissue when the ratio of FPKM values was at least four-fold. Transcripts with FPKM values lower than 1 in all tissues are not shown.

expressed between male and female chemosensory appendages in *R. suturalis*.

DISCUSSION

In this study, we have identified the first chemosensory genes of an aquatic member of the megadiverse order Coleoptera, the diving beetle *Rhantus suturalis*. Members of six chemosensory gene families (OBPs, CSPs, NPC2 proteins, and receptors of the OR, GR, and IR families) were annotated, incorporated into phylogenetic analyses, and their relative expression levels were estimated, based on RNA-seq data from antennae and palps of adult males and females as well as wing muscle (i.e., in total, 5 samples). The gene repertoires recovered from these transcriptomic data are necessarily partial in the absence of a reference genome for *R. suturalis*, and because only the main cephalic sensory organs at a single life stage were processed. This

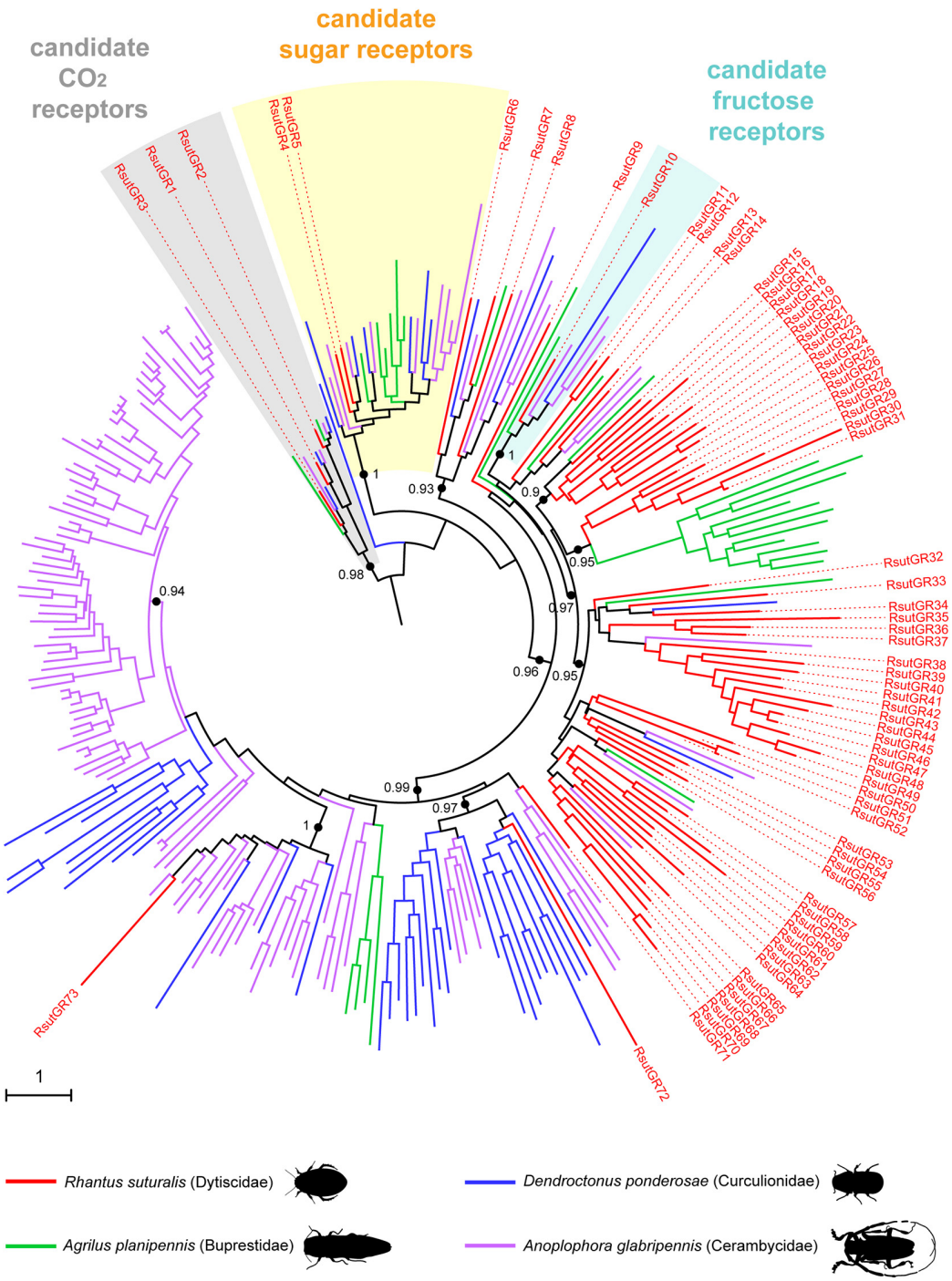

FIGURE 5 | Maximum-likelihood phylogeny of coleopteran GRs. The tree was built from an alignment of amino acid sequences of *R. suturalis* GRs (RsutGR, available in **Supplementary Table 1**) with sequences annotated from the genomes of three coleopteran species belonging to distinct superfamilies. Putative functions were assigned to several clades as in Andersson et al. (2019). Black dots indicate deep nodes highly supported by the likelihood-ratio test (aLRT > 0.9), with the corresponding value. The tree was rooted using the CO_2 receptor clade as an outgroup. The scale bar indicates the expected number of amino acid substitutions per site.

limitation has to be kept in mind, especially for comparisons of gene family sizes with respect to other Coleoptera or insect species for which genes have been annotated based on complete genomes. However, given our relatively deep sequencing strategy, and results from the BUSCO analysis (only 4.1% of genes missing), we can surmise that the contents obtained here

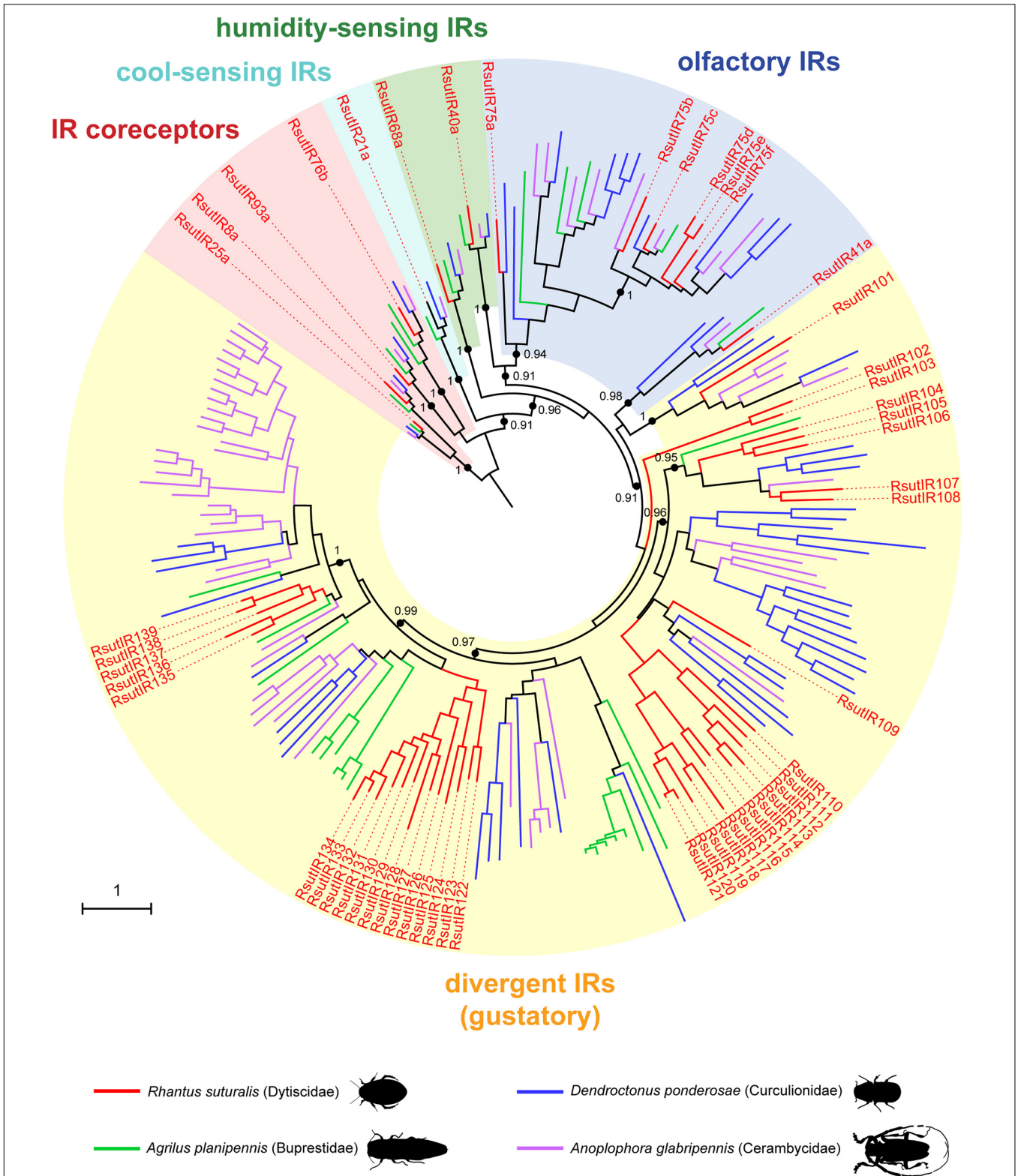

FIGURE 6 | Maximum-likelihood phylogeny of coleopteran IRs. The tree was built from an alignment of amino acid sequences of *R. suturalis* IRs (RsutIR, available in **Supplementary Table 1**) with sequences annotated from the genomes of three coleopteran species belonging to distinct superfamilies. Putative functions were assigned to the different IR clades based on previous works on *Drosophila*. Black dots indicate deep nodes highly supported by the likelihood-ratio test (aLRT > 0.9), with the corresponding value. The tree was rooted using the IR25a/IR8a clade as an outgroup. The scale bar indicates the expected number of amino acid substitutions per site.

unigene	BlastP UniProt best hit	Species	UniProt ID	E-value	male ant.	female ant.	male palps	female palps	wing muscle
cytochrome P450									
contig_108714	Uncharacterized protein	*Cryptotermes secundus*	A0A2J7R4Q4	9.67e-25					
contig_108707	Cytochrome P450 CYP6BQ22	*Dastarcus helophoroides*	M9TLT7	1.80e-28					
contig_160300	Uncharacterized protein	*Photinus pyralis*	A0A1Y1K8X2	2.18e-21					
contig_159068	Cytochrome P450 12H1	*Tribolium castaneum*	D6WP99	4.51e-32					
contig_48897	Putative cytochrome P450 6a14	*Cryptotermes secundus*	A0A2J7R4Q0	3.09e-114					
contig_160295	Cytochrome P450 monooxygenase CYP6BQ35	*Tenebrio molitor*	A0A0K1YWE6	2.08e-134					
contig_114244	Cytochrome P450 301B1	*Tribolium castaneum*	D2A103	0.0					
contig_72829	Uncharacterized protein	*Photinus pyralis*	A0A1Y1K2A7	2.42e-156					
contig_156394	Uncharacterized protein	*Glossina brevipalpis*	A0A1A9W296	5.75e-105					
contig_187886	Putative cytochrome P450 6a14	*Cryptotermes secundus*	A0A2J7R4P8	1.18e-166					
contig_243870	Putative cytochrome P450 6a13	*Zootermopsis nevadensis*	A0A067RK29	5.69e-45					
contig_162369	Cytochrome P450 6BQ13	*Tribolium castaneum*	D6WB59	5.68e-151					
contig_64438	Uncharacterized protein	*Photinus pyralis*	A0A1Y1KCR3	7.02e-141					
carboxylesterase family									
contig_356304	Carboxylic ester hydrolase	*Harmonia axyridis*	E2QC35	9.83e-26					
contig_317173	Carboxylic ester hydrolase	*Liposcelis bostrychophila*	B6D7H0	2.95e-77					
contig_317168	Carboxylic ester hydrolase	*Harmonia axyridis*	E2QC35	4.75e-26					
contig_263675	Carboxylic ester hydrolase	*Tenebrio molitor*	A0A0K1YW40	4.65e-61					
contig_263672	Carboxylic ester hydrolase	*Leptinotarsa decemlineata*	A0A0A7ENL4	1.51e-12					
contig_333714	Carboxylic ester hydrolase	*Leptinotarsa decemlineata*	A0A0A7EQM9	2.28e-114					
contig_333705	Carboxylic ester hydrolase	*Dendroctonus ponderosae*	U4UFS6	2.76e-39					
lipase									
contig_129485	Ves G 1 allergen precursor, putative	*Pediculus humanus*	E0VE21	5.32e-22					
contig_284122	PNLIPRP2_0 protein	*Fopius arisanus*	A0A0C9RCJ3	1.76e-37					
contig_329878	Uncharacterized protein	*Apis mellifera*	A0A088AM94	9.99e-30					
short chain dehydrogenase									
contig_346543	Dehydrogenase/reductase SDR family	*Zootermopsis nevadensis*	A0A067QVP7	3.33e-23					
contig_346530	farnesol dehydrogenase-like	*Agrilus planipennis*	A0A1W4WN42	3.02e-18					
contig_275110	Uncharacterized protein	*Apis mellifera*	H9K5X4	7.48e-24					
contig_60810	Uncharacterized protein	*Graphocephala atropunctata*	A0A1B6M980	7.65e-15					
contig_275114	Uncharacterized protein	*Musca domestica*	A0A1I8MBG2	3.00e-16					
contig_275113	Dehydrogenase/reductase SDR family	*Blattella germanica*	A0A2P8XJR8	5.23e-31					
contig_216873	Uncharacterized protein	*Tribolium castaneum*	D7ELE5	4.37e-77					
contig_111444	Zgc:56585	*Astyanax mexicanus*	W5JZW2	9.34e-46					
contig_319907	Uncharacterized protein	*Glossina pallidipes*	A0A1A9Z224	5.35e-16					
contig_335408	Dehydrogenase	*Oryctes borbonicus*	A0A0T6B086	1.19e-66					
contig_102214	15-hydroxyprostaglandin dehydrogenase	*Blattella germanica*	A0A140KPR8	3.49e-62					
contig_111449	15-hydroxyprostaglandin dehydrogenase	*Lingula unguis*	A0A1S3IUI5	4.49e-15					
contig_281499	farnesol dehydrogenase-like	*Agrilus planipennis*	A0A1W4WZ38	1.04e-74					
contig_281511	Dehydrogenase/reductase SDR family	*Zootermopsis nevadensis*	A0A067QVP7	2.76e-29					
contig_62298	Uncharacterized protein	*Tribolium castaneum*	D7ELE5	3.48e-74					
contig_33141	Uncharacterized protein	*Photinus pyralis*	A0A1Y1LK58	1.20e-102					
contig_235599	Uncharacterized protein	*Photinus pyralis*	A0A1Y1N8N2	3.94e-85					
UDP-glucuronosyl and UDP-glucosyl transferase									
contig_114450	UDP-glucuronosyltransferase	*Agrilus planipennis*	A0A1W4XHV4	2.35e-106					
contig_47820	UDP-glucuronosyltransferase	*Photinus pyralis*	A0A1Y1N2G3	2.21e-99					
contig_81795	UDP-glucuronosyltransferase	*Pediculus humanus*	E0VMM1	2.05e-75					
CD36 family									
contig_93159	scavenger receptor class B member 1-like	*Agrilus planipennis*	A0A1W4WGT2	7.27e-38					
contig_93152	scavenger receptor class B member 1-like	*Agrilus planipennis*	A0A1W4WGT2	1.22e-78					
contig_21684	sensory neuron membrane protein 1-like	*Agrilus planipennis*	A0A1W4WII7	4.60e-35					
contig_200289	Sensory neuron membrane protein 2	*Holotrichia parallela*	A0A2P1ERQ2	9.63e-130					
contig_161176	Sensory neuron membrane protein SNMP-1	*Tenebrio molitor*	A0A0C5DMP8	1.21e-168					
contig_284635	scavenger receptor class B member 1-like	*Agrilus planipennis*	A0A1W4WTV9	0.0					
amiloride-sensitive sodium channel									
contig_162809	Sodium channel protein Nach	*Zootermopsis nevadensis*	A0A067QUS0	1.70e-46					
contig_349482	Pickpocket protein 11-like Protein	*Tribolium castaneum*	D7EHU3	7.98e-69					
contig_37680	Sodium channel protein Nach-like Protein	*Tribolium castaneum*	A0A139WEM9	2.70e-121					

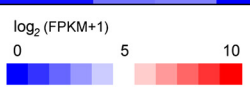

FIGURE 7 | List of *R. suturalis* unigenes specifically expressed in antennae and palps and potentially involved in chemical senses. Protein families were defined following Pfam protein domains found by hmmscan (see detail in section "Materials and Methods"). Color coding in the heatmap is based on $\log_2$(FPKM + 1) values; raw FPKM values are available in **Supplementary Table 2**.

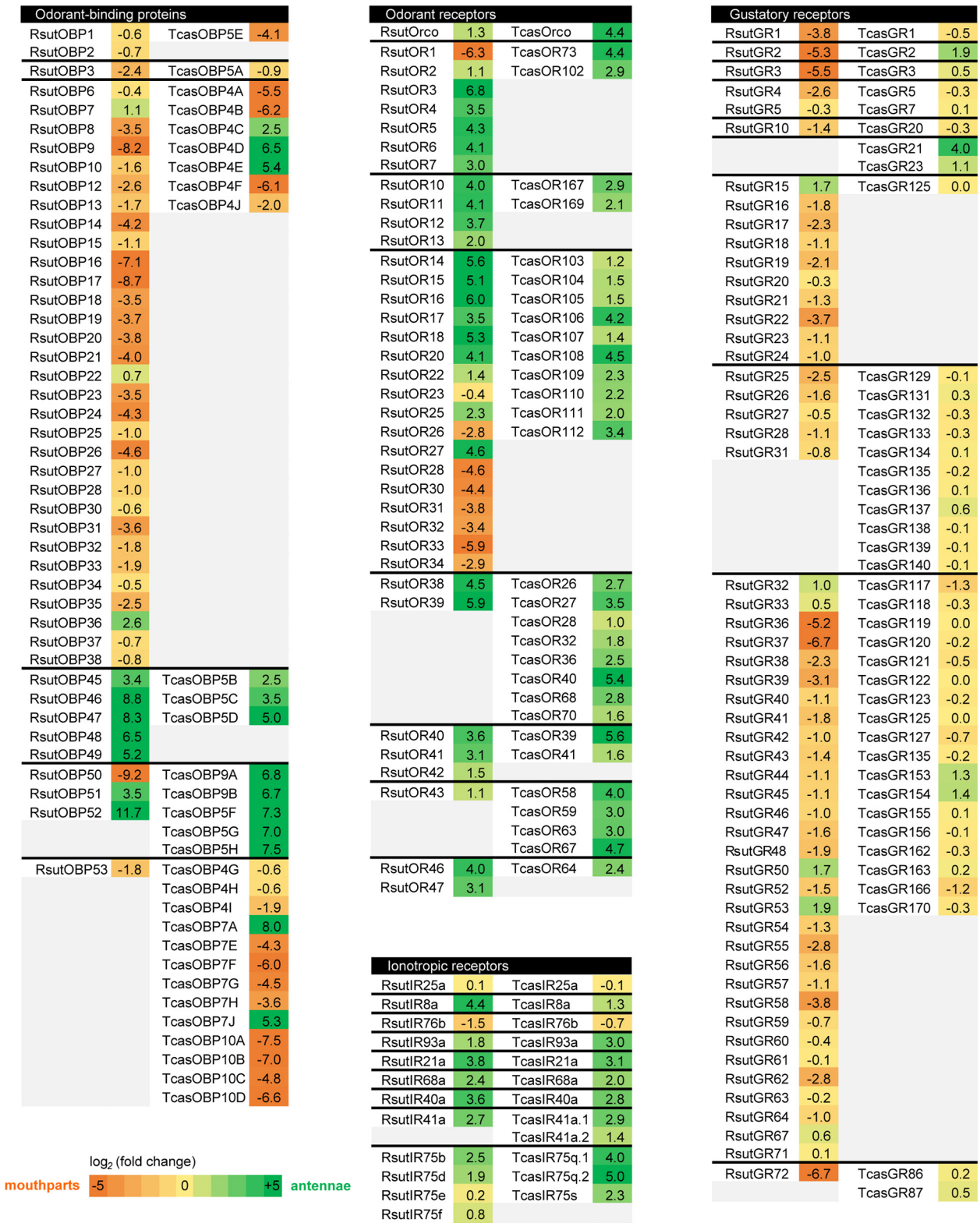

FIGURE 8 | Heatmaps comparing chemosensory gene expression bias in antennae vs. maxillary and labial palps (or entire mouthparts) in *R. suturalis* and *T. castaneum*. Fold changes were calculated based on FPKM values (*R. suturalis*) or RPKM values (*T. castaneum*, Dippel et al., 2014, 2016), and color coding shows $\log_2$(fold change). For each family, genes are separated by groups of orthology.

for the different gene families should not be considerably far from exhaustiveness.

Our *R. suturalis* transcriptome yielded chemosensory gene repertoires whose sizes are generally within the range of what has been previously obtained for terrestrial Coleoptera species, but in the lower, middle or upper range depending of the gene families. *R. suturalis* has more OBPs (53 retrieved in this study) than the Curculionidae *D. ponderosae* and the Buprestidae *A. planipennis* but this number is of the same order of magnitude as in genomes of the Tenebrionidae *T. castaneum*, the Chrysomelidae *Leptinotarsa decemlineata* and the Cerambycidae *A. glabripennis* (Andersson et al., 2019). Among other roles, OBPs are thought to mediate transport of hydrophobic odorants through the sensillar lymph to the olfactory neuron dendrites (Brito et al., 2016; Rihani et al., 2021). Numerous functional studies carried out on Coleoptera (mainly in Scarabaeidae and Chrysomelidae) have identified OBPs capable of binding hydrophobic molecules, whether they are aggregation pheromones or plant volatiles (see Mitchell and Andersson, 2021 for a review). Interestingly, this holds true for the Dytiscidae *Cybister japonicus*, where a classical OBP and a Minus-C OBP have shown a good affinity for a monoterpene (citral) and for a phenypropanoid (coniferyl aldehayde), respectively (Song et al., 2016). The fact that *R. suturalis* has a rich complement of OBPs suggests that detection of hydrophobic odorants is important in the chemical ecology of this aquatic beetle. Moreover, the OBP phylogeny revealed an impressive gene expansion in the lineage leading to the Dytiscidae, with a clade of classical OBPs exclusively containing 35 *R. suturalis* paralogs (**Figure 1**). In its breadth, this expansion is unparalleled in any other beetle species investigated. This may reflect peculiar functional requirements on OBPs in aquatic Adephaga, perhaps due to the low concentration of hydrophobic molecules in water. However, no OBP has been described yet in terrestrial Adephaga and it remains to be determined whether this large OBP expansion is linked with the transition to an aquatic lifestyle or not.

Contrary to OBPs, the *R. suturalis* transcriptome contained a low number of NPC2 proteins and CSPs, which are other soluble proteins potentially involved in chemical senses in insects. Whereas the low number of NPC2 proteins is on a par with previous observations in Coleoptera (Pelosi et al., 2014), the number of four CSPs identified in *R. suturalis* is far below what has been found in beetle genomes (11–20; Andersson et al., 2019). Together with the fact that their expression was generally not restricted to antennae and palps, this suggests that these two protein families may not play a major role in chemoreception in *R. suturalis* and aquatic Adephaga. Potential roles of CSPs in detoxification, immunity or secretion of defensive compounds have already been proposed in *T. castaneum* (Contreras et al., 2013; Li et al., 2013).

For chemoreceptors, the number of 48 ORs identified in *R. suturalis* stands in the lower range for Coleoptera (32–258; Mitchell et al., 2020). Like OBPs, ORs are supposed to be functionally involved in the detection of hydrophobic odorants. It is interesting to observe that this diving beetle species does not present an impoverished complement of ORs with respect to the terrestrial Adephaga beetle *C. scrutator* (51 ORs). Of the Coleoptera species whose OR genes have been annotated from complete genomes, only the Hydroscaphidae *Hydroscapha redfordi* has less ORs than *R. suturalis* (32). Hydroscaphidae are aquatic members of suborder Myxophaga, feeding on algae. The OR phylogeny revealed a possible Dytiscidae-specific OR gene expansion within clade 3, harboring 9 *R. suturalis* paralogs. Lineage-specific expansions have previously been found in this OR clade of unknown function for species with different lifestyles and feeding habits, namely the carnivorous burying beetle *N. vespilloides* and the phytophagous long-horned beetle *A. glabripennis*. Nevertheless, *R. suturalis* together with its terrestrial Adephaga cousin *C. scrutator* underwent much less massive gene expansions in the whole OR phylogeny than observed in Polyphaga species for which data are available (Mitchell et al., 2020).

We identified 73 GRs from the *R. suturalis* transcriptome, mostly expressed in palps. This number is higher than the GR gene repertoire of specialized phytophagous Polyphaga but much lower than in polyphagous species of Polyphaga (e.g., *T. castaneum*: 219; *A. glabripennis*: 190; Andersson et al., 2019). Our phylogenetic analysis showed that GR expansions identified in *R. suturalis* did not occur in the same lineages as those observed in *A. glabripennis* and *D. ponderosae*. However, the current lack of GR functional characterization in Coleoptera hampers any discussion on the potential role of these expansions. As insect GRs are mostly known for their ability to bind water-soluble non-volatile semiochemicals, it is tempting to speculate that a larger GR complement would increase underwater chemodetection capabilities but this has yet to be demonstrated. Moreover, the lack of GR gene identification in any other Adephaga makes it impossible to determine which of these GR expansions observed in *R. suturalis* are possibly specific to water beetles. The same is true for lineage-specific expansions involving *R. suturalis* divergent IR genes mostly expressed in palps.

In the absence of biological replicates, we could not carry out a differential expression analysis able to capture subtle differences of expression between samples, notably between males and females. Anyway, no *R. suturalis* chemosensory gene showed any hint of sex-biased expression, a quite puzzling result in light of the recent experimental demonstration that in this species, males are attracted underwater by a sex pheromone (of unknown composition) emitted by females (Herbst et al., 2011). With the exception of *D. melanogaster*, all pheromone receptors identified so far in insects belong to the OR family (see Fleischer and Krieger, 2021 for a comprehensive review). Even though receptors to aggregation pheromone components (such as those identified in Coleoptera) can be expressed at similar levels in both sexes, sex pheromone receptors generally present a strong sex-biased expression, as commonly observed in male moths (Bastin-Heline et al., 2019). It is of course conceivable that in diving beetles, receptors allowing males to detect the female pheromone belong to a receptor family other than those examined in this study. Other possible explanations could be that a pheromone receptor is indeed present among the candidate receptors that we characterized but is not differentially expressed between the sexes, or that expression of the pheromone receptor is seasonal and was not different between sexes at the period of

the year in which we sampled the specimens used in this study (between April and October 2017).

One of the most interesting features of the *R. suturalis* chemosensory gene repertoires in the face of the terrestrial-aquatic ecological transition is that possible dytiscid-specific gene expansions were systematically associated with expression in palps. The comparison of chemosensory gene expression levels in antennae vs. palps (or entire mouthparts) in *R. suturalis* and *T. castaneum* (the only terrestrial beetle for which comparable data are available; Dippel et al., 2014, 2016) reveals that preferential expression sites clearly differ for several orthologous genes or gene groups, with a striking recurrent tendency across chemosensory gene families: expression tends to be shifted to the palps in *R. suturalis* (**Figure 8**). Amongst ORs, *R. suturalis* has several palp-specific ORs whose *T. castaneum* orthologs are antennae-specific: RsutOR1 (member of clade 1) and all but one members of the *R. suturalis*-specific expansion in clade 3 (RsutOR26–34). Concerning GRs, most *T. castaneum* genes are expressed similarly in antennae and mouthparts and some are antennae-biased, whereas the vast majority of RsutGRs seem to be more expressed in palps, notably the highly expressed RsutGR36, 37 and 72, as well as the candidate CO_2 receptors RsutGR1-3. Contrary to the other two chemoreceptor families, expression of IRs is highly similar in *T. castaneum* and *R. suturalis*, yet this comparison does not include divergent IRs, which have not been analyzed in *T. castaneum* (Dippel et al., 2016). Finally, several *T. castaneum* orthologs of the large *R. suturalis* OBP expansion (RsutOBP5–39) are also expressed at a higher level in palps, but RsutOBP50 (belonging to the ABP II clade) is highly expressed in palps whereas its *T. castaneum* orthologs are all antenna-specific.

CONCLUSION

In conclusion, our transcriptome analysis of chemosensory organs in a diving beetle has revealed chemosensory gene repertoires that are rather similar to those of terrestrial Coleoptera. This seems at odds with prevailing views on the impact of physical constraints on chemical communication underwater. However, we did identify several occurrences of genes highly diversified in *R. suturalis*, notably in the OBP and OR families, which are associated to the detection of hydrophobic odorants in insects. Moreover, these diversified genes were generally highly expressed in palps. A shift of expression of OBPs and ORs from the antennae to the palps may have significant implication with respect to the ecological transition from terrestrial to aquatic life. Indeed, hydrophobic molecules are much less diffusive in water than in the air, leading to the idea that the same proteins that mediate long-range chemodetection (i.e., olfaction in the classic sense) in aerial animals may be involved in short-range or contact chemodetection (gustation, or taste, in the classic sense) in aquatic animals (Mollo et al., 2014). Further expression analyses in Adephaga with various lifestyles associated with functional studies of their chemosensory genes will be necessary to verify whether part of the molecular toolkit functioning in the antennae of terrestrial beetles has indeed diversified and has been re-allocated to the palps in aquatic beetles, to allow short-range perception of hydrophobic molecules.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI (accession: PRJNA762278).

AUTHOR CONTRIBUTIONS

NM, MJ, TC, and MM conceived the study. MJ and MM performed insect collection and sample preparation. YJ prepared RNA-seq libraries and performed sequencing. NM, TC, EP, CM, and EJ-J performed transcriptome assembly, data analysis, and phylogenetic reconstructions. NM and MM drafted the manuscript, with input from all authors. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the Alliance Sorbonne Université (EMERGENCE project Beetle-EvoChem) and the French National Research Agency (ANR-19-CE02-003-01 grant). This work has benefited from the expertise of the High-throughput Sequencing Platform of I2BC.

ACKNOWLEDGMENTS

We thank Paula Graça and Latifa Belkessa (ISYEB, Sorbonne Université) for technical help and Thierry Jaffredo (IBPS, Sorbonne Université) for laboratory facilities.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.773915/full#supplementary-material

REFERENCES

Andersson, M. N., Keeling, C. I., and Mitchell, R. F. (2019). Genomic content of chemosensory genes correlates with host range in wood-boring beetles (*Dendroctonus ponderosae*, *Agrilus planipennis*, and *Anoplophora glabripennis*). *BMC Genomics* 20:690. doi: 10.1186/s12864-019-6054-x

Anisimova, M., and Gascuel, O. (2006). Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst. Biol.* 55, 539–552. doi: 10.1080/10635150600755453

Baker, G. T. (2001). Distribution patterns and morphology of sensilla on the apical segment of the antennae and palpi of Hydradephaga (*Coleoptera: adephaga*). *Microsc. Res. Tech.* 55, 330–338. doi: 10.1002/jemt.1181

Bastin-Heline, L., de Fouchier, A., Cao, S., Koutroumpa, F., Caballero-Vidal, G., Robakiewicz, S., et al. (2019). A novel lineage of candidate pheromone receptors for sex communication in moths. *ELife* 8:e49826. doi: 10.7554/eLife.49826

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Brand, P., Robertson, H. M., Lin, W., Pothula, R., Klingeman, W. E., Jurat-Fuentes, J. L., et al. (2018). The origin of the odorant receptor gene family in insects. *ELife* 7:e38340. doi: 10.7554/eLife.38340

Brito, N. F., Moreira, M. F., and Melo, A. C. (2016). A look inside odorant-binding proteins in insect chemoreception. *J. Insect Physiol.* 95, 51–65. doi: 10.1016/j.jinsphys.2016.09.008

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176

Chen, Z., Wang, Q., and Wang, Z. (2010). The amiloride-sensitive epithelial Na^+ channel PPK28 is essential for drosophila gustatory water reception. *J. Neurosci.* 30, 6247–6252. doi: 10.1523/JNEUROSCI.0627-10.2010

Chertemps, T., and Maïbèche, M. (2021). "Odor degrading enzymes and signal termination," in *Insect Pheromone Biochemistry and Molecular Biology*, 2nd Edn, eds G. J. Blomqvist and R. G. Vogt (Amsterdam: Elsevier), 619–644. doi: 10.1016/B978-0-12-819628-1.00019-5

Cock, P. J., Chilton, J. M., Gruning, B., Johnson, J. E., and Soranzo, N. (2015). NCBI BLAST+ integrated into Galaxy. *Gigascience* 4:39. doi: 10.1186/s13742-015-0080-7

Contreras, E., Rausell, C., and Real, M. D. (2013). Proteome response of *Tribolium castaneum* larvae to *Bacillus thuringiensis* toxin producing strains. *PLoS One* 8:e55330. doi: 10.1371/journal.pone.0055330

Croset, V., Rytz, R., Cummins, S. F., Budd, A., Brawand, D., Kaessmann, H., et al. (2010). Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet.* 6:e1001064. doi: 10.1371/journal.pgen.1001064

Culler, L. E., Ohba, S., and Crumrine, P. (2014). "Predator-Prey Interactions of Dytiscids," in *Ecology, Systematics, and the Natural History of Predaceous Diving Beetles (Coleoptera: Dytiscidae)*, ed. D. Yee (Dordrecht: Springer), 363–386. doi: 10.1007/978-94-017-9109-0_8

Dijkstra, K. D., Monaghan, M. T., and Pauls, S. U. (2014). Freshwater biodiversity and aquatic insect diversification. *Annu. Rev. Entomol.* 59, 143–163. doi: 10.1146/annurev-ento-011613-161958

Dippel, S., Kollmann, M., Oberhofer, G., Montino, A., Knoll, C., Krala, M., et al. (2016). Morphological and transcriptomic analysis of a beetle chemosensory system reveals a gnathal olfactory center. *BMC Biol.* 14:90. doi: 10.1186/s12915-016-0304-z

Dippel, S., Oberhofer, G., Kahnt, J., Gerischer, L., Opitz, L., Schachtner, J., et al. (2014). Tissue-specific transcriptomics, chromosomal localization, and phylogeny of chemosensory and odorant binding proteins from the red flour beetle *Tribolium castaneum* reveal subgroup specificities for olfaction or more general functions. *BMC Genomics* 15:1141. doi: 10.1186/1471-2164-15-1141

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113. doi: 10.1186/1471-2105-5-113

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., et al. (2019). The Pfam protein families database in 2019. *Nucleic Acids Res.* 47, D427–D432. doi: 10.1093/nar/gky995

Eyun, S. I., Soh, H. Y., Posavi, M., Munro, J. B., Hughes, D. S. T., Murali, S. C., et al. (2017). Evolutionary history of chemosensory-related gene families across the arthropoda. *Mol. Biol. Evol.* 34, 1838–1862. doi: 10.1093/molbev/msx147

Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39(Web Server issue), W29–W37. doi: 10.1093/nar/gkr367

Fleischer, J., and Krieger, J. (2021). "Molecular mechanisms of pheromone detection," in *Insect Pheromone Biochemistry and Molecular Biology* -, 2nd Edn, eds G. J. Blomqvist and R. G. Vogt (Amsterdam: Elsevier), 355–413. doi: 10.1016/B978-0-12-819628-1.00012-2

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565

Giordano, G., Carbone, M., Ciavatta, M. L., Silvano, E., Gavagnin, M., Garson, M. J., et al. (2017). Volatile secondary metabolites as aposematic olfactory signals and defensive weapons in aquatic environments. *Proc. Natl. Acad Sci. U.S.A.* 114, 3451–3456. doi: 10.1073/pnas.1614655114

Gouy, M., Guindon, S., and Gascuel, O. (2010). SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* 27, 221–224. doi: 10.1093/molbev/msp259

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512. doi: 10.1038/nprot.2013.084

Hansson, B. S., and Stensmyr, M. C. (2011). Evolution of insect olfaction. *Neuron* 72, 698–711. doi: 10.1016/j.neuron.2011.11.003

Herbst, C., Baier, B., Tolasch, T., and Steidle, J. L. M. (2011). Demonstration of sex pheromones in the predaceous diving beetle *Rhantus suturalis* (MacLeay 1825) (Dytiscidae). *Chemoecology* 21, 19–23. doi: 10.1007/s00049-010-0061-3

Hodgson, E. S. (1953). A study of chemoreception in aqueous and gas phases. *Biol. Bull.* 105, 115–127. doi: 10.2307/1538560

Isono, K., and Morita, H. (2010). Molecular and cellular designs of insect taste receptor system. *Front. Cell Neurosci.* 4:20. doi: 10.3389/fncel.2010.00020

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., and Madden, T. L. (2008). NCBI BLAST: a better web interface. *Nucleic Acids Res.* 36(Web Server issue), W5–W9. doi: 10.1093/nar/gkn201

Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* 20, 1160–1166. doi: 10.1093/bib/bbx108

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Koh, T. W., He, Z., Gorur-Shandilya, S., Menuz, K., Larter, N. K., Stewart, S., et al. (2014). The Drosophila IR20a clade of ionotropic receptors are candidate taste and pheromone receptors. *Neuron* 83, 850–865. doi: 10.1016/j.neuron.2014.07.012

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315

Leal, W. S. (2013). Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu. Rev. Entomol.* 58, 373–391. doi: 10.1146/annurev-ento-120811-153635

Lefort, V., Longueville, J. E., and Gascuel, O. (2017). SMS: Smart Model Selection in PhyML. *Mol. Biol. Evol.* 34, 2422–2424. doi: 10.1093/molbev/msx149

Li, J., Lehmann, S., Weissbecker, B., Ojeda Naharros, I., Schutz, S., Joop, G., et al. (2013). Odoriferous defensive stink gland transcriptome to identify novel genes necessary for quinone synthesis in the red flour beetle, *Tribolium castaneum*. *PLoS Genet.* 9:e1003596. doi: 10.1371/journal.pgen.1003596

Liu, C., Pitts, R. J., Bohbot, J. D., Jones, P. L., Wang, G., and Zwiebel, L. J. (2010). Distinct olfactory signaling mechanisms in the malaria vector mosquito *Anopheles gambiae*. *PLoS Biol.* 8:467. doi: 10.1371/journal.pbio.1000467

Liu, T., Wang, Y., Tian, Y., Zhang, J., Zhao, J., and Guo, A. (2020). The receptor channel formed by ppk25, ppk29 and ppk23 can sense the *Drosophila* female pheromone 7,11-heptacosadiene. *Genes Brain Behav.* 19:e12529. doi: 10.1111/gbb.12529

Matthews, B. J., Younger, M. A., and Vosshall, L. B. (2019). The ion channel ppk301 controls freshwater egg-laying in the mosquito *Aedes aegypti*. *ELife* 8:e43963. doi: 10.7554/eLife.43963.001

McKenna, D. D., Scully, E. D., Pauchet, Y., Hoover, K., Kirsch, R., Geib, S. M., et al. (2016). Genome of the Asian longhorned beetle (*Anoplophora glabripennis*), a globally significant invasive species, reveals key functional and evolutionary

innovations at the beetle-plant interface. *Genome Biol.* 17:227. doi: 10.1186/s13059-016-1088-8

Mitchell, R. F., and Andersson, M. N. (2021). "Olfactory genomics of the Coleoptera," in *Insect Pheromone Biochemistry and Molecular Biology*-, 2nd Edn, eds G. J. Blomqvist and R. G. Vogt (Amsterdam: Elsevier), 547–589. doi: 10.1016/B978-0-12-819628-1.00017-1

Mitchell, R. F., Schneider, T. M., Schwartz, A. M., Andersson, M. N., and McKenna, D. D. (2020). The diversity and evolution of odorant receptors in beetles (*Coleoptera*). *Insect Mol. Biol.* 29, 77–91. doi: 10.1111/imb.12611

Mollo, E., Fontana, A., Roussis, V., Polese, G., Amodeo, P., and Ghiselin, M. T. (2014). Sensing marine biomolecules: smell, taste, and the evolutionary transition from aquatic to terrestrial life. *Front. Chem.* 2:92. doi: 10.3389/fchem.2014.00092

Montell, C. (2009). A taste of the *Drosophila* gustatory receptors. *Curr. Opin. Neurobiol.* 19, 345–353. doi: 10.1016/j.conb.2009.07.001

Ng, R., Salem, S. S., Wu, S. T., Wu, M., Lin, H. H., Shepherd, A. K., et al. (2019). Amplification of *Drosophila* olfactory responses by a DEG/ENaC channel. *Neuron* 104, 947–959. doi: 10.1016/j.neuron.2019.08.041

Nielsen, H. (2017). Predicting secretory proteins with SignalP. *Methods Mol. Biol.* 1611, 59–73. doi: 10.1007/978-1-4939-7015-5_6

Pelosi, P., Iovinella, I., Felicioli, A., and Dani, F. R. (2014). Soluble proteins of chemical communication: an overview across arthropods. *Front. Physiol.* 5:320. doi: 10.3389/fphys.2014.00320

Prieto-Godino, L. L., Rytz, R., Cruchet, S., Bargeton, B., Abuin, L., Silbering, A. F., et al. (2017). Evolution of Acid-Sensing Olfactory Circuits in Drosophilids. *Neuron* 93, 661–666. doi: 10.1016/j.neuron.2016.12.024

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the european molecular biology open software suite. *Trends Genet.* 16, 276–277. doi: 10.1016/s0168-9525(00)02024-2

Rihani, K., Ferveur, J. F., and Briand, L. (2021). The 40-year mystery of insect odorant-binding proteins. *Biomolecules* 11:509. doi: 10.3390/biom11040509

Rimal, S., and Lee, Y. (2018). The multidimensional ionotropic receptors of *Drosophila melanogaster*. *Insect Mol. Biol.* 27, 1–7. doi: 10.1111/imb.12347

Robertson, H. M. (2019). Molecular evolution of the major arthropod chemoreceptor gene families. *Annu. Rev. Entomol.* 64, 227–242. doi: 10.1146/annurev-ento-020117-043322

Ruel, D. M., Yakir, E., and Bohbot, J. D. (2018). Supersensitive odorant receptor underscores pleiotropic roles of indoles in mosquito ecology. *Front. Cell Neurosci.* 12:533. doi: 10.3389/fncel.2018.00533

Sanchez-Alcaniz, J. A., Silbering, A. F., Croset, V., Zappia, G., Sivasubramaniam, A. K., Abuin, L., et al. (2018). An expression atlas of variant ionotropic glutamate receptors identifies a molecular basis of carbonation sensing. *Nat. Commun.* 9:4252. doi: 10.1038/s41467-018-06453-1

Shan, S., Wang, S. N., Song, X., Khashaveh, A., Lu, Z. Y., Dhiloo, K. H., et al. (2019). Antennal ionotropic receptors IR64a1 and IR64a2 of the parasitoid wasp *Microplitis mediator* (*Hymenoptera: Braconidate*) collaboratively perceive habitat and host cues. *Insect Biochem. Mol. Biol.* 114:103204. doi: 10.1016/j.ibmb.2019.103204

Short, A. E. Z. (2018). Systematics of aquatic beetles (*Coleoptera*): current state and future directions. *Syst. Entomol.* 43, 1–18. doi: 10.1111/syen.12270

Silbering, A. F., Rytz, R., Grosjean, Y., Abuin, L., Ramdya, P., Jefferis, G. S., et al. (2011). Complementary function and integrated wiring of the evolutionarily distinct *Drosophila* olfactory subsystems. *J. Neurosci.* 31, 13357–13375. doi: 10.1523/JNEUROSCI.2360-11.2011

Simao, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351

Simon, P. (2003). Q-Gene: processing quantitative real-time RT-PCR data. *Bioinformatics* 19, 1439–1440. doi: 10.1093/bioinformatics/btg157

Song, L. M., Jiang, X., Wang, X. M., Li, J. D., Zhu, F., Tu, X. B., et al. (2016). Male tarsi specific odorant-binding proteins in the diving beetle *Cybister japonicus* Sharp. *Sci. Rep.* 6:31848. doi: 10.1038/srep31848

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578. doi: 10.1038/nprot.2012.016

van Giesen, L., and Garrity, P. A. (2017). More than meets the IR: the expanding roles of variant Ionotropic Glutamate Receptors in sensing odor, taste, temperature and moisture. *F1000Res* 6:1753. doi: 10.12688/f1000research.12013.1

Wicher, D., and Miazzi, F. (2021). "Insect odorant receptors: Function and regulation," in *Insect Pheromone Biochemistry and Molecular Biology*, 2nd Edn, eds G. J. Blomqvist and R. G. Vogt (Amsterdam: Elsevier), 415–433. doi: 10.1016/B978-0-12-819628-1.00013-4

Xia, Y., Wang, G., Buscariollo, D., Pitts, R. J., Wenger, H., and Zwiebel, L. J. (2008). The molecular and cellular basis of olfactory-driven behavior in *Anopheles gambiae* larvae. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6433–6438. doi: 10.1073/pnas.0801007105

Xu, W. (2020). How do moth and butterfly taste? Molecular basis of gustatory receptors in *Lepidoptera*. *Insect Sci.* 27, 1148–1157. doi: 10.1111/1744-7917.12718

Yohe, L. R., and Brand, P. (2018). Evolutionary ecology of chemosensation and its role in sensory drive. *Curr. Zool.* 64, 525–533. doi: 10.1093/cz/zoy048

Zhang, J., Bisch-Knaden, S., Fandino, R. A., Yan, S., Obiero, G. F., Grosse-Wilde, E., et al. (2019). The olfactory coreceptor IR8a governs larval feces-mediated competition avoidance in a hawkmoth. *Proc. Natl. Acad. Sci U.S.A.* 116, 21828–21833. doi: 10.1073/pnas.1913485116

Zhang, X. X., Yang, B., Sun, D. D., Guo, M. B., Zhang, J., and Wang, G. R. (2021). Ionotropic receptor 8a is involved in the attraction of *Helicoverpa armigera* to acetic acid. *Insect Sci.* doi: 10.1111/1744-7917.12962 [Epub ahead of print].

Zhao, Y. J., Li, G. C., Zhu, J. Y., and Liu, N. Y. (2020). Genome-based analysis reveals a novel SNMP group of the Coleoptera and chemosensory receptors in *Rhaphuma horsfieldi*. *Genomics* 112, 2713–2728. doi: 10.1016/j.ygeno.2020.03.005

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Montagné, Jager, Chertemps, Persyn, Jaszczyszyn, Meslin, Jacquin-Joly and Manuel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.